

INFORMATIKA

Bobřík učí informatiku

5. díl – Informace a jejich reprezentace

DANIEL LESSNER – JIŘÍ VANÍČEK

Matematicko-fyzikální fakulta, UK Praha

Pedagogická fakulta, Jihočeská univerzita v Českých Budějovicích

Slovo počítač je sice odvozeno od slova počítat, dnes už se ale na zpracovávaná data díváme spíš jako na informace, než jako na čísla. Tento rozdíl je nicméně skutečně jen v úhlu pohledu.

Ve škole se žáci typicky dozvědí, že všechna data v počítači jsou uložena pomocí nul a jedniček, že základní jednotkou informace je bit, jaké jsou násobné jednotky a jak s tím vším souvisí problematika velikosti souborů a přenosové rychlosti. Naučí se počítat kapacitu daného počtu bitů, tedy proč např. bajt (osmice bitů) umožňuje rozlišit 256 možností. Tato oblast informatiky úzce souvisí s matematikou. Poměrně jednoduchou strukturou, vhodnou pro první seznámení s kombinatorikou, jsou právě slova dané délky z dané abecedy (tedy variace s opakováním).

Méně často se žáci dozvědí, co to vlastně informace je, protože jde o dosti komplexní pojem. Žáci obvykle pracují s intuitivní představou a hlubší vztah mezi informacemi a např. množstvím dat, která jsou třeba k jejich uchování, jim zůstává skryt.

Jeden ze způsobů seznámení s podstatou informace se opírá o model vylučování možností. Ukažme si příklad: k narozeninám očekávám bicykl, knihu, tablet nebo deskovou hru. Všechny možnosti považuji za stejně možné. Pohledem na zabalený dárek v krabici pak získám nějakou informaci, např. je zjevné, že to nebude kolo. Potězkáním vytuším, jestli to bude kniha, zachrastění by odhalilo deskovou hru. Něco se dozvědět, tedy získat informaci, zde znamená vyloučit jednu nebo několik možností. Naopak, po-

kud nám někdo něco říká, ale nevyklučuje tím žádné možnosti (např. dárek byl koupen v e-shopu), žádné informace nezískáváme. Čím více možností se nabízí, tím více informace potřebujeme k určení té správné.

Mají-li žáci aspoň základní představu o pravděpodobnosti (jako poměru počtu sledovaných a počtu všech jevů), lze pojem informace zobecnit. Nevím-li nic o tom, jak bude zítra, znamená to jinými slovy, že považuji všechny možnosti za stejně pravděpodobné. Pokud ale třeba na základě své zkušenosti, místa na Zemi, roční doby či počasí posledních dnů odhaduji na 70 % slunečno, o počasí něco vím. Kolik přesně toho vím, závisí právě na rozložení pravděpodobnosti jednotlivých variant. Když se něco nového dozvím, rozložení pravděpodobností se změní. Rozdíl mezi novým a původním rozložením odpovídá množství získané informace.

Například pokud házím ideálně spravedlivou hrací kostkou, o výsledku netuším nic a pravděpodobnost všech šesti výsledků je $1/6$. Pokud si všimnu, že někdo podstrčil kostku „cinknutou“, počítám s vyšší pravděpodobností pádu např. šestky (na úkor pravděpodobností ostatních). Odhalení falešnosti kostky tedy o výsledku hodu nese určité množství informace. Až se kostka dokutálí, bude pravděpodobnost skutečného výsledku 1 a ostatních 0. Hodem kostky všechnu neurčitost o výsledku odstraníme a všechnu informaci získáme. Více o této problematice (včetně vztahu informace a entropie) uvádí učebnice Informatika pro každého [1].

Uvedené pojetí informace souvisí s užitím pravděpodobnostního počtu. Pravděpodobnostně totiž modelujeme situace, o nichž máme málo informací. Uvažujeme v nich více než jednu možnost zároveň a pravděpodobnostní ohodnocení jednotlivých možností vlastně odpovídá tomu, co o situaci víme. Nijak přitom nezáleží na tom, jestli jsou v pozadí nějaké náhodné jevy, nebo tomu pozadí prostě jen nerozumíme. Více o informacích v souvislosti s pravděpodobností píše Bastl v článku Teorie informace [2].

V následujících úlohách chceme ukázat, jak informace omezuje počet možností, jak k jejímu uchování používáme bity a bajty nebo jak lze informace různým způsobem reprezentovat.

Uhodni tvar

Kategorie Junior, autorka Daniela Bezáková.

Zadání

Hraješ hru proti počítači. Na obrazovce je 9 obrázců (obrázek). Vybereš si jeden z nich, ale počítači neoznámíš, který. Pak se počítač začne ptát a

ty musíš pravdivě odpovídat ANO nebo NE.



Počítač smí dávat tyto otázky:

Je tvůj obrazec červený?

Je tvůj obrazec modrý?

Je tvůj obrazec žlutý?

Je tvůj obrazec kruh?

Je tvůj obrazec čtverec?

Je tvůj obrazec trojúhelník?

Úkolem počítače je uhodnout tvůj obrazec co nejdříve. Ty víš, že počítač je naprogramován hrát tuto hru nejlepším možným způsobem.

Kolik otázek musí počítač položit, aby jistě uhodl tvůj obrazec?

Co má tato úloha společného s informatikou

Jde o úlohu, při níž hledáme postup, jak omezit možnosti, tedy jak se získáním nějakých informací dostat k jedné správné možnosti. K vyřešení úlohy je potřeba objevit vhodný algoritmus omezování možností a spočítat jeho „časovou náročnost“. Potom může takovéto otázky pokládat místo člověka stroj a budeme si jisti, kolik otázek (a tím i času) bude potřeba na zjištění toho, co chceme vědět.

Zdůvodnění správné odpovědi

V úloze je třeba třídit podle dvou kritérií: tvar (3 možnosti) a barva (3 možnosti). Vezmeme např. tvar: k rozlišení, který ze tří tvarů je správný, potřebujeme dvě otázky. Totéž platí pro barvu. Jsou tedy potřeba $2+2 = 4$ otázky.

Kdyby hádal člověk, mohl by se podle situace rozhodnout a některé otázky vynechat. V některých případech by mu stačily i 2 otázky (kdyby se náhodou napoprvé strefil do barvy i tvaru). S jistotou ovšem nemůžeme říci, že mu vždy budou stačit 2 nebo 3 otázky, aby obrazec uhodl. Proto nemůže být správná odpověď 2 ani 3.

Počítač je naprogramován, pracuje s algoritmem na nalezení hledaného tvaru. Algoritmu budou také stačit nanejvýš 4 otázky. A stejně jako člověk může na získané informace reagovat. Pokud např. na první otázku, týkající se tvaru, dostane odpověď ANO, není třeba se na tvar dotazovat podruhé.

Levák nebo pravák?

Kategorie Senior, autor Jiří Vaníček.

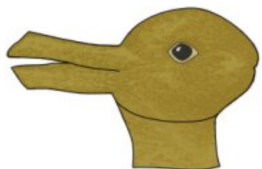
Zadání

Psychologové prováděli u žáků testy laterality (tedy jestli je levák nebo pravák). Dali jim tři úkoly a výsledky zapisovali do počítače.

1. Sepněte ruce. (Psychologové do počítače zapsali, jestli žák měl nahoře levý nebo pravý palec.)



2. Podívejte se na obrázek. Rychle řekněte, jaké zvíře na něm vidíte. (Zapsali, zda žák viděl hlavu zajíce nebo kachny.)



3. Zatleskejte. (Zapsali, která ruka byla při tleskání nahoře, levá nebo pravá.)



Kolik nejméně paměti počítače je potřeba k zaznamenání odpovědi jednoho žáka?

- A) 3 bity B) 8 bitů C) 3 bajty D) 8 bajtů

Co má tato úloha společného s informatikou

K vyřešení úlohy je třeba znát základní jednotky informace a jejich význam. Řešitel musí zjistit, kolik je možných různých výsledků, a pomocí

kolika bitů či bajtů lze takový počet možností rozlišit. Podobně se ptá informatik např. při navrhování databáze zboží: od počtu různých položek se odvíjí délka identifikačního čísla. Identifikačním číslem je i rodné číslo. Nám v Česku slouží dobře, ale např. v Číně by naše rodné číslo nevyhovovalo proto, že nemůže nést dostatek informace k identifikaci jedince. Číňanů se denně narodí víc než 10 000, jenže v našem rodném čísle jsou k rozlišení dětí, narozených v daný den, za lomítkem k dispozici pouze 4 číslice.

Při navrhování sítě Internet se informatici zamýšleli, kolik se k němu nanejvýš bude připojovat počítačů, a podle toho volili délku IP adresy, tedy jednoznačného identifikátoru zařízení v síti. V té době nepředpokládali, že by různých zařízení mohlo být víc než čtyři miliardy, a zvolili proto čtyřbajtové adresy. Vznikající síť podcenili, takže se v současné době rozšiřuje standard adres pro jistotu rovnou šestnáctibajtových. Takové úvahy se opírají o stejné principy jako řešení této úlohy. Více o standardech IP adres na <http://cs.wikipedia.org/wiki/IPv6>.

Zdůvodnění správné odpovědi

Jeden bit (neboli dvojková číslice, tedy 0 nebo 1) je nejmenší jednotka informace. O bitech nejčastěji hovoříme v souvislosti s pamětí počítače. Do paměti o velikosti jednoho bitu lze uložit informaci určující jednu ze dvou možností (např. prší/neprší, správná odpověď/špatná odpověď, levák/pravák, pravdivé/nepravdivé).

Do dvou bitů uchováme informaci o jedné ze čtyř možností. Např. pokud budeme rozlišovat leváctví a pohlaví dotazovaných, mohou nastat čtyři možnosti (chlapec levák, chlapec pravák, dívka levák, dívka pravák) a ty zaznamenáme čtyřmi kombinacemi nul a jedniček (00, 01, 10, 11) – první bit zde označuje pohlaví, druhý laterální. Každý úkon testování laterality má dvě možné odpovědi (levý nebo pravý palec, kachna nebo zajíc, levá nebo pravá ruka je nahoře). Odpověď se dvěma možnostmi se vejde do paměti velikosti 1 bitu, když označíme např. 0 – zajíc, 1 – kachna a obdobně odpovědi na ostatní otázky. Na každou odpověď je potřeba 1 bit, na 3 odpovědi tedy postačí stačí 3 bity. Správná odpověď je A).

Bajt je jednotka informace, která obsahuje 8 bitů (možnost B) a může tak určit jednu z 256 možností (viz <http://cs.wikipedia.org/wiki/Bajt>). Tři bajty (možnost C) mohou rozlišit $256 \cdot 256 \cdot 256 = 256^3 = 16\,777\,216$ možností. Osm bajtů (možnost D) jich rozliší ještě mnohem víc (256^8), je jasné, že to správné řešení nebude.

Barevná hloubka

Kategorie Junior, autor Peter Tomcsányi.

Zadání

Nekomprimovaný bitmapový obrázek má rozměry 40 krát 50 pixelů a zabírá v paměti přesně 1000 bajtů.

Kolik nejvíce různých barev může používat?

- A) více než 16 000 B) 256 C) 128 D) 16

Co má tato úloha společného s informatikou

Tato úloha ukazuje další praktické využití výše popsaného vztahu mezi množstvím informace (či dat) a počtem možností. Čím věrnější barvy požadujeme, tím více jich musíme rozlišit, a tím více paměti zabere výsledný obrázek. Aby se vůbec dalo mluvit o obrázku, potřebujeme barvy aspoň dvě, např. černou a bílou, které si zakódujeme např. 0 a 1. Černobílý obrázek tedy potřebuje pro uchování informace o 1 pixelu 1 bit.

Dnešní zařízení běžně zobrazují miliony různých barevných odstínů a dostávají se tak ke hranici toho, co dovede rozlišit lidské oko a mozek. Potom ovšem potřebují k popsání barvy v jednom pixelu mnoho bitů. Čím více paměti je třeba na popsání barev pixelů, tím větší barevnou hloubku má daný obrázek. Např. pro běžně používané zobrazení True Color, v němž se používá asi 16,8 milionu barevných odstínů, je k jejich rozlišení třeba 24 bitů, tedy 3 bajty na každý pixel.

Více viz https://cs.wikipedia.org/wiki/Barevná_hloubka.

Zdůvodnění správné odpovědi

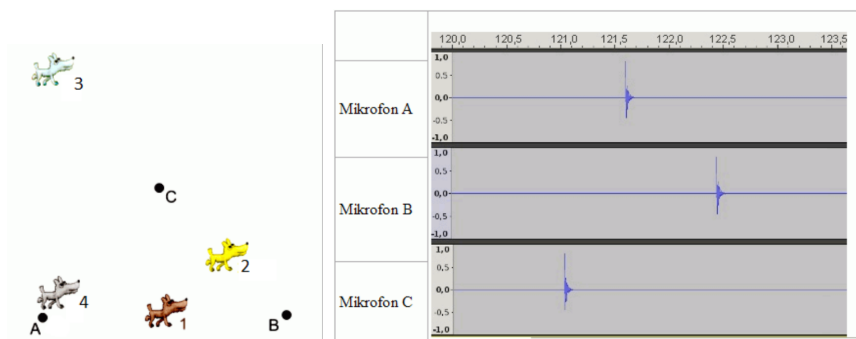
Obrázek má $40 \cdot 50$, tedy 2000 pixelů. Jeden bajt je 8 bitů, 1000 bajtů je 8000 bitů. Na jeden pixel tak připadají $8000/2000 = 4$ bity. Čtyřmi bity lze rozlišit nanejvýš $2^4 = 16$ barevných kombinací. Obrázek může mít nanejvýš 16 barev.

Akustická inteligence

Kategorie Junior, autor Heno Ivanov.

Zadání

Na podlaze stojí tři mikrofony (obrázek, černé tečky ukazují jejich umístění). Najednou jeden ze psů zaštěkal. Všechny tři mikrofony zachytily tento zvuk tak, jak ukazuje graf.



Který ze čtyř psů na obrázku zaštěkal?

1 – hnědý 2 – žlutý 3 – bělavý 4 – šedivý

Co má tato úloha společného s informatikou

K určení polohy v prostoru jsme zvyklí používat souřadnice. Není to však jediná možnost, zejména v situaci, kdy potřebujeme lokalizovat objekt, který něco vysílá (či přijímá), a máme k dispozici přijímače (nebo vysílače) na známých polohách. Přitom může jít o štěkající psy a mikrofony, o mobilní telefony a bezdrátově připojené počítače nebo třeba o GPS přijímače a satelity.

Poloha je pak určena vzdálenostmi od daných konkrétních bodů a může být reprezentována např. časovými údaji o zaznamenané události. Jinými slovy, tyto časové údaje o tom, kdy signály dorazily k přijímači, nesou stejnou informaci jako souřadnice.

Zdůvodnění správné odpovědi

Jak je vidět z grafu, mikrofon C zachytil zaštěkání psa nejdříve. Bělavý pes 3 je jediný pes, k němuž je mikrofon C nejbližší ze všech mikrofonů. Hnědý pes 1 má stejně daleko ke všem třem mikrofonům, žlutý pes 2 má stejně daleko k mikrofonu B a C, šedivý pes 4 má nejbližší k mikrofonu A. Správně je bělavý pes 3.

Zmenšování fotografií

Kategorie Junior, autor Jiří Vaníček.

Zadání

Pepík měl fotografii sochy v souboru v počítači (viz obrázek).



Nejprve obrázek v grafickém editoru zmenšil na 25 % a uložil ho. Poté nový soubor otevřel a zvětšil jej na 400 %. Jak potom obrázek vypadal?



A)



B)



C)



D)

Co má tato úloha společného s informatikou

Tato úloha se týká rastrových (bodových, bitmapových) obrázků a jejich škálování. Informace zachycená na digitální fotografii (na rozdíl třeba od vektorové grafiky) je uchována v jednotlivých pixelech (obrazových bodech). Jejich počet přímo souvisí s množstvím detailů, které lze na obrázku rozlišit.

Změnou velikosti obrázku se změní počet pixelů. Barvy nových pixelů přitom vychází z barev pixelů na odpovídajících místech původního obrázku. Zmenšovací algoritmy se snaží co možná nejlépe uchovat původní obrázek. Pokud např. původní velký obrázek obsahuje černou čáru o tloušťce jediného pixelu, je třeba dobře promyslet, co s ní: máme ji nechat splynout s pozadím, čímž možná obětujeme podstatný detail obrázku, nebo ji máme zachovat, čímž ji ale v novém obrázku nepoměrně zvýrazníme?

Zmenšením různých větších a vzájemně rovnocenných obrázků může vzniknout tentýž obrázek. Úloha zpětného zvětšování tudíž nemá jednoznačnou odpověď. Různé algoritmy preferují různé zvětšené obrázky podle účelu, kterému mají sloužit. Vylepšováním algoritmů pro škálování obrázků se zabývají inmatičtí zaměřeni na počítačovou grafiku.

Zdůvodnění správné odpovědi

Správná odpověď je B). Obrázek se nejprve zmenší na 25 % a poté zvětší na 400 % (čtyřikrát se zmenší a pak se čtyřikrát zvětší). Jeho rozměry se tedy nezmění. Odpovědi A) a D) jsou nesprávné.

Po zmenšení obrázku každému novému pixelu odpovídá $4 \times 4 = 16$ pixelů z původního obrázku. Pixel má ale jen jednu barvu, není možné v něm uchovat 16 barev. Při zmenšení rastrového obrazu se proto část obrazové informace ztrácí. Ani zvětšením na původní (nebo libovolnou jinou) velikost tuto původní informaci nelze obnovit, není odkud ji vzít.

Konečný obrázek nemůže vypadat stejně jako původní. Místo původních detailů bude rozmazaný, „kostičkováný“ či jinak poškozený, což odpovídá odpovědi B). Odpověď C) by byla správně, kdybychom obrázek nejprve zvětšovali a pak zmenšovali.

Literatura

- [1] *Lessner, D.*: Učebnice informatiky pro každého. [online] MFF UK, 2015. Dostupné z <http://ksvi.mff.cuni.cz/ucebnice/index.php/Učebnice/Informace>.
- [2] *Bartl, E.*: Teorie informace. MFI, roč. 24, č. 3, str. 219–228.
- [3] Předchozí díly seriálu Bobřík učí informatiku najdete v MFI roč. 22, č. 5 (úvodní díl), roč. 23, č. 2 (procházení grafů), roč. 23, č. 4 (algoritmické úlohy), roč. 24, č. 5 (použití logiky v informatice).